PB86-208691

Practical Considerations for Approximating Relative Risk by the Standardized Mortality Ratio

North Carolina Univ. at Chapel Hill

Prepared for

Health Effects Research Lab.
Research Triangle Park, NC

1981

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before comp

1. REPORT NO.	2.	PB86-208691
EPA/600/J-81/699		

4. TITLE AND SUBTITLE
PRACTICAL CONSIDERATIONS FOR APPROXIMATING RELATIVE RISK BY THE STANDARDIZED MORTALITY RATIO
(Journal Version)

5. REPORT DATE
1981

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
M. J. Symons* and J. D. Taulbee

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
*Symons, Dept. of Biostatistics and Occupational Health Studies Group UNC,NC
Taulbee, Procter and Gamble Co., Cincinnati, Ohio

10. PROGRAM ELEMENT NO.

11. CONTRACT/GRANT NO.

12. SPONSORING AGENCY NAME AND ADDRESS
HEALTH EFFECTS RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NC 27711

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE
EPA/600/11

15. SUPPLEMENTARY NOTES
John P. Creason, BD/HERL/

REF: Journal of Occupational Medicine, 23(6):413-416, June 1981

16. ABSTRACT

The standardized mortality ratio is a widely used and often criticized summary statistic for studies of occupational mortality. In this paper the authors discuss practical conditions under which relative risk can reasonably be approximated by the standardized mortality ratio. When the true relative risk is greater than 100%, the standardized mortality ratio overestimates relative risk no matter how small the mortality rates or how short the age bands utilized in the analysis. However, when the excessive mortality is consistent across the age bands, the standardized mortality ratio can usefully approximate relative risk for some applications, such as those involving the site-specific cancers, providing the age bands employed are not too large.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	6
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE

EPA Form 2220-1 (Rev. 4-77)

REPRODUCED BY
NATIONAL TECHNICAL INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This article documents research that was conducted under the auspices of the U.S. Environmental Protection Agency (USEPA). As an arm of the government, the USEPA has a paid up, non-exclusive, irrevocable, worldwide license to reproduce, prepare derivative works, and distribute to the public all of the material contained herein, and to grant others acting on its behalf the right to do so. The USEPA grants to the National Technical Information Service, Department of Commerce, the right to act as its agent in the distribution of this material.

EPA/600/J-81/699
JOURNAL ARTICLE

Practical Considerations for Approximating Relative Risk by the Standardized Mortality Ratio

M. J. Symons, Ph.D., and J. D. Taulbee, Ph.D.

The standardized mortality ratio is a widely used and often criticized summary statistic for studies of occupational mortality. In this paper the authors discuss practical conditions under which relative risk can reasonably be approximated by the standardized mortality ratio. When the true relative risk is greater than 100%, the standardized mortality ratio overestimates relative risk no matter how small the mortality rates or how short the age bands utilized in the analysis. However, when the excessive mortality is consistent across the age bands, the standardized mortality ratio can usefully approximate relative risk for some applications, such as those involving site-specific cancers, providing the age bands employed are not too large.

The standardized mortality ratio (SMR) has been generally criticized and its value as a useful summary statistic for studies of occupational mortality has been seriously questioned in two articles previously appearing in the *Journal of Occupational Medicine*.[1,2] The utility of the SMR as an approximation to relative risk (RR) was one of several characteristics considered. In this paper it is pointed out that the SMR can be a useful approximation to RR under certain conditions. Specifically it is suggested that the SMR reasonably approximates relative risk when (1) the age-specific rates in the comparison population for the cause(s) of interest are no larger than about 100 per 10,000 workers per year with ten-year age bands. As this death rate of 0.0100 corresponds roughly to the all-malignancies mortality rate among U.S. white males about 1970, applications involving site-specific cancers are candidates for good approximation of RR by the SMR; (2) the age bands utilized in the analysis are not too broad. The mathematical aspects indicate that the products of the rate and the length of the age bands are factors that determine the adequacy of the approximation. As would be expected, shorter age bands are needed when the age-specific mortality rates are large, and longer ones can be utilized when the death rates are small. Fewer deaths are observed when the mortality rates are small and, unfortunately, in practice there may be a tendency to employ longer age bands just to avoid zero frequencies of deaths. Extending the guideline for 10-year age bands, it is noted that with 20-year age bands, the mortality rates should be no larger than about 50 per 10,000 workers per year for the approximation to be reasonable. With five-year age bands these rates may be as large as 200 per 10,000 workers per year; and (3) the age-specific mortality rates for the study and comparison populations are in approximately constant ratio across the age bands. One cannot expect this condition to hold strictly in most applications, but it can be approached by allowing an adequate latency period for a cancer or a chronic condition to develop and by not including older age bands after most of the disease has manifested itself. In short, the age range for the analysis should not begin at too early an age nor extend too far into older ages; otherwise the true relative risk may be underestimated.

From the Department of Biostatistics and Occupational Health Studies Group, University of North Carolina, Chapel Hill, NC 27514 (Dr Symons), and the Procter and Gamble Company, Miami Valley Laboratories, Cincinnati, OH 45247 (Dr Taulbee)

This research was completed while Dr Symons was on leave with an Intergovernmental Personnel Act appointment with the Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle, NC 27711. A major portion was completed while Dr Taulbee was on the faculty of the Department of Biostatistics and Occupational Health Studies Group, University of North Carolina

Definitions and Background

The work force of interest in this study is composed of a single sex-race group so that the only factor to be adjusted is age. Table 1 contains the notation for the available data in one age group for the study (exposed) and comparison (unexposed) populations. As is often the case,[3,4,5,6] the data for the study population are from a cohort of workers and those for the comparison popula-

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

Table 1. — Notation for the Mortality Experience for One Age Band in Two Populations.				
Population	No. of Deaths Due to Cause of Interest	Person-Years or Mid-Year Population	Mortality Rate	Probability* of Dying
Study (exposed)	d	n	r = d/n	q
Comparison (unexposed)	D	N	R = D/N	Q

*Full definition is in the text

tion are from (cross-sectional) vital statistics and census surveys for the same sex-race group in the U.S. general population.

If the comparison population is also available as a cohort of individuals, the Mantel-Haenszel[7] weighted average of the age-specific odds ratios is an appropriate alternative to the SMR, as Redmond and Breslin[8] have commented. Cornfield[9] showed that the odds ratio is an approximation to relative risk when the probability of disease is small, so that a similar relation between the odds ratio and relative risk exists, as can be shown[10] for the SMR and for relative risk.

The ratio of the number of deaths in the exposed population to the number expected without the exposure(s) is the mortality ratio (MR) in the specified age group. Specifically, it is noted from Table 1 that

$$MR = d/Rn = r/R. \quad (1)$$

The relative risk for the same age group is the ratio of the probability of an individual in the exposed population who is alive at the beginning of the age interval dying during the age band to the same probability for an individual in the unexposed population. It is also observed from Table 1 that

$$RR = q/Q. \quad (2)$$

Here the relative risk is a concept involving the life chances of two individuals while the mortality ratio in (1) compares the death rates for the two populations. These are two related,[11] but different, ideas. Since relative risk pertains to individuals rather than to populations, as do death rates, the interpretation of relative risk is quite meaningful and an object for estimation.

The standardized mortality ratio is given by

$$SMR = \frac{O}{E} = \sum d/\sum Rn = \sum rn/\sum Rn, \quad (3)$$

the summations being over the age groups. The SMR then is the ratio of the observed (O) to the expected (E) total number of deaths, where the expected number is computed by applying the schedule of age-specific mortality rates (R) in the comparison population to the distribution of person-years (n) in the study population. The expression at the right in (3) shows the SMR as the ratio of a weighted average of the schedule of the age-specific rates (r) in the study population to one of the age-specific rates (R) in the comparison population. Unfortunately the age-specific weights are the numbers of person-years of experience (n) in the study population, so that comparison of SMRs is a questionable practice. This point has long been known[12,13] and has been reemphasized by others.[14,15] In fact, various unrealistic examples are easily constructed[1] to illustrate this point but differences also occur in practice. Wong[2] further discusses the problem and the literature. On the other hand, if the distributions of person-years in the two study populations are similar and if the same comparison population is used for each SMR, the ratio of two such SMRs may not be misleading, as can be seen from (3).

Few alternatives to the SMR are available. Miettinen[14] suggested using direct adjusted rates, but when the age-specific mortality rates and numbers of person-years of experience are both small, the age-specific rates (r) in the study population may fluctuate widely if one death more or less is observed in the corresponding age band. When the age-specific mortality rates in the study population are stable, the direct adjustment procedure leads to an analogue[14] of the SMR. Enterline[16] employed age-specific death rates in an occupational mortality study, but due to the lack of stability of the age-specific rates (r), the practice of using direct adjustment of mortality rates has generally not been followed in occupational mortality studies.

Despite inherent weaknesses, the standardized mortality ratio is *the* summary statistic in most occupational mortality studies. In support of this usage the essence of the approximation of relative risk by the SMR is outlined under certain conditions. If the mortality rates for the exposed and unexposed populations are in constant ratio, specifically

$$r/R = \theta \text{ or } r = \theta R, \quad (4)$$

for all age groups, then from (3) the result is

$$SMR = \sum(\theta R)n/\sum Rn = \theta. \quad (5)$$

Using a linear approximation to an exponential function, it can be shown[10] that the relative risk is approximately equal to the same constant θ. This approximation is reasonable when the age-specific death rates are small, the age bands are not too large, and the mortality rates for the study and comparison populations are in constant ratio across all the age bands.

Numerical Illustrations

The approximation of relative risk by the SMR depends upon three factors. These are: (1) the magnitude of the mortality rate (R) in the comparison population; (2) the size of the true relative risk (θ) which is presumed constant over age; and (3) the length of age bands (t) utilized in the analysis. The range of occupational mortality applications where the SMR reasonably approximates relative risk is quickly grasped with a few numerical illustrations.

First, the effect of the magnitude of the mortality rates has to be considered in the comparison population. This is determined largely by the cause(s) of death and the age range under study. In Table 2 the mortality rates for all malignant neoplasms are given by 10-year age bands for the 1971 U.S. white male population,[17] aged 35 to 74

Table 2. — Mortality Ratios by 10-Year Age Bands for All Malignant Neoplasms (ICDA Codes: 140-209) for 1971 U.S. White Males and a Fictitious Work Force with a Relative Risk of 200%.

Age Group	Probability of Dying		Relative Risk (%)	Mortality Rate		Mortality Ratio (%)
	1971 U.S. White Males*	Fictitious Work Force		1971 U.S. White Males[17]	Fictitious Work Force*	
35-44	0.004898	0.009796	200	0.000491	0.000984	200.4
45-54	0.016778	0.033556	200	0.001692	0.003413	201.7
55-64	0.047905	0.095810	200	0.004909	0.010072	205.2
65-74	0.096392	0.192784	200	0.010136	0.021416	211.3

*Derived using a fundamental relationship between the probability of dying and the mortality rate that is given by Chiang[11] and detailed for this application by Symons and Taulbee[10]; t = 10 years

years, along with the experience of a fictitious work force, devised[10,11] so that the age-specific relative risks are 200%. This means that, in each 10-year age band, the probability of death from any malignancy for an individual in the fictitious work force who is alive at the beginning of the age band is twice that of a U.S. white male of the same age in 1971. Depending upon the age distribution of the person-years of experience in the fictitious work force, it is deduced (3) that the SMR for Table 2 will be between 200.4% and 211.3%. Therefore the SMR would overestimate the RR by no more than 5.65%.

With smaller mortality rates the approximation is much more accurate. In Table 3 the mortality rates for lung cancers are presented in parallel fashion to Table 2, except the relative risk is presumed to be 300%. The resultant mortality ratios range from 300.7% to 311.5%. An overestimation of RR by the SMR in this situation would be no larger than 3.85%.

The effect of the true relative risk and the length of the age bands is considered in Table 4. The mortality rate of 100 per 10,000 workers per year was chosen for illustration and as a practical upper limit for many occupational mortality applications of the SMR since it is approximately the death rate for *all* malignant neoplasms for U.S. white males, aged 65 to 74 years in 1971. The overestimation of RR by the SMR, as a percentage excess, is given in the right most column of Table 4 for each of three age bands: 5 years, 10 years, and 20 years. Within these age bands the true relative risk is presumed to be 200%, 300%, and 400%, respectively. The SMR overestimates the RR by varying amounts that can be seen to increase steadily with the true relative risk and the length of the age bands.

These numerical illustrations can be summarized by noting several things. First, for large age-specific death rates the SMR approximation to relative risk may be a gross overestimate. Gaffey[1] provides an example in his Table 3 for *all* causes of mortality. However, for some causes of death of interest in occupational mortality the SMR can provide a very satisfactory approximation to RR, as with the lung cancer example in Table 3 of this study. Second, when the true relative risk is very large, greater than 400%, for example, the SMR will be even larger, as can be seen from Table 4. Finally, with the use of large age bands, the resulting SMR will overestimate the true relative risk. However, as can also be noted in Table 4, by the use of short enough age bands all of the aforementioned overestimations can be controlled to a satisfactory degree for many practical occupational mortality applications, such as site-specific cancers.

Discussion and Conclusion

The standardized mortality ratio can approximate relative risk when the mortality rates are not too large, the age bands are not too long, and the age range for the analysis is not too broad. Several remarks about these circumstances and the nature of this approximation are noteworthy.

First, the choice of the intervals of age to be employed in the analysis is an important design consideration for each application, since the approximation presumes a constant relative risk across the age range. For example, with many site-specific malignancies, it may be necessary to allow a sufficient latent period in order to detect an elevated relative risk. Furthermore, inclusion of age bands at the upper age range where the relative risk is not elevated can also have an effect of decreasing the sensitivity of the analysis. Dement et al[18] describe a study of respiratory cancer among chrysotile asbestos textile workers where sufficient latency is allowed and where also most of the lung cancer seems to have manifested itself by about age 65. Enterline[19] covers a number of

Table 3. — Mortality Ratios by 10-Year Age Bands for Lung Cancer (ICDA Codes: 160-163) for 1971 U.S. White Males and a Fictitious Work Force with a Relative Risk of 300%.

Age Group	Probability of Dying		Relative Risk (%)	Mortality Rate		Mortality Ratio (%)
	1971 U.S. White Males*	Fictitious Work Force		1971 U.S. White Males[17]	Fictitious Work Force*	
35-44	0.001469	0.004407	300	0.000147	0.000442	300.7
45-54	0.006668	0.020004	300	0.000669	0.002021	302.1
55-64	0.019448	0.058344	300	0.001964	0.006012	306.1
65-74	0.035398	0.106194	300	0.003604	0.011227	311.5

*Derived using a fundamental relationship between the probability of dying and the mortality rate that is given by Chiang[11] and detailed for this application by Symons and Taulbee[10]; t = 10 years

Table 4. — Mortality Ratio as a Function of Relative Risk and Length of Age Interval with a Mortality Rate of 100 per 10,000 Workers per Year.

Length of Age Band	Probability of Dying		Relative Risk RR%	Mortality Rate		Mortality Ratio MR%	Percentage Excess (MR - RR) x 100 / RR
	Comparison Population*	Study Population		Comparison Population	Study Population*		
5 years	0.048770	0.097540	200	0.010	0.020526	205.3	2.6
5 years	0.048770	0.146310	300	0.010	0.031637	316.4	5.5
5 years	0.048770	0.195080	400	0.010	0.043402	434.0	8.5
10 years	0.095163	0.190326	200	0.010	0.021112	211.1	5.6
10 years	0.095163	0.285489	300	0.010	0.033616	336.2	12.1
10 years	0.095163	0.380652	400	0.010	0.047909	479.1	19.8
20 years	0.181269	0.362538	200	0.010	0.022513	225.1	12.6
20 years	0.181269	0.543807	300	0.010	0.039242	392.4	30.8
20 years	0.181269	0.725076	400	0.010	0.064563	645.6	61.4

*Derived using a fundamental relationship between the probability of dying and the mortality rate that is given by Chiang[11] and detailed for this application by Symons and Taulbee[10]. t = 5, 10, or 20 years as indicated

study design points and provides an excellent discussion of this basic study design issue.

Second, whenever the true relative risk is greater than unity (100%), the approximation by the SMR is anti-conservative however small the death rates may be. The magnitude of this overestimation increases with the value of the magnitude of the mortality rates in the comparison population, the true relative risk, and the length of the age bands. In particular, with 10-year age bands and age-specific mortality rates no larger than 100 per 10,000 workers per year, the SMR overestimates the RR by no more than 6%, 13%, or 20% when the true relative risk is 200%, 300%, or 400%, respectively. However, the amount of overestimation can be controlled by utilizing shorter length age bands, especially when the true relative risk is expected to be significantly large or the age-specific rates in the comparison population become fairly large at the upper age range of the analysis.

Finally, it is emphasized that the relative risk is unlikely to be constant with age. Nevertheless it is important to restrict the age range where the relative risk is excessive so that an average *elevated* risk will be estimated by the SMR. If the death experience is sufficient in the cohort, the mortality ratio for each five-year age band, for example, can be calculated and the pattern of excess risk with age can be reported.

The standardization of the SMR provides a *single* summary statistic from several age-specific mortality rates which can estimate an average elevated RR. From the weighted average form displayed in equation (3), the SMR will lie between the smallest and largest of the mortality ratios of the study to the comparison population age-specific rates — e.g., when these age-specific mortality ratios are all equal, the SMR will be this common value. Difficulties in the SMR are due largely to this weighted average construction. Chiazze's comment[19] that there is no single substitute for a schedule of age-specific rates is very relevant, because such a schedule provides the only way to confirm the validity of the fundamental assumption of consistency in the elevated risk across age which is needed for the SMR to reasonably approximate RR. This is an often repeated recommendation in the occupational mortality study methods literature[19,20] and should be followed whenever the data are sufficient to permit it.

References

1. Gaffey WR. A critique of the standardized mortality ratio. *J Occup Med* **18**:157-160, 1976.
2. Wong O. Further criticisms on epidemiological methodology in occupational studies. *J Occup Med* **19**:220-222, 1977.
3. Enterline PE. Mortality among asbestos products workers in the United States. *Ann NY Acad Sci* **132**:156-165, 1975.
4. Selikoff IJ, Churg J, and Hammond EC. Asbestos exposure and neoplasia. *JAMA* **188**:22-26, 1964.
5. Enterline PE, DeCoufle P, and Henderson V. Respiratory cancer in relation to occupational exposures among retired asbestos workers. *Br J Ind Med* **30**:162-166, 1973.
6. Andjelkovic D, Taulbee JD, and Symons MJ. Mortality experience of a cohort of rubber workers, 1964-1973. *J Occup Med* **18**:387-394, 1976.
7. Mantel N and Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* **22**:719-748, 1959.
8. Redmond CK and Breslin PP. Comparison of methods for assessing occupational hazards. *J Occup Med* **17**:313-317, 1975.
9. Cornfield J. A method of estimating comparative rates from clinical data applications to cancer of the lung, breast and cervix. *J Natl Cancer Inst* **11**:1269-1275, 1951.
10. Symons MJ and Taulbee JD. Standardized mortality ratio as an approximation to relative risk. Institute of Statistics Mimeo Series No. 1294, July, 1980.
11. Chiang CL. A stochastic study of the life table and its applications. I. Probability distributions of the biometric functions. *Biometrics* **16**:618-635, 1960.
12. Yule GU. On some points relating to vital statistics, more especially statistics of occupational mortality. *J Royal Stat Soc* **94**:1-84, 1934.
13. Hill AB. Principles of Medical Statistics, Sixth Edition. Oxford University Press, New York, N.Y., 1959.
14. Miettinen OS. Standardization of risk ratios. *Am J Epidemiol* **96**:383-385, 1972.
15. Chiazze L. Problems of study design and interpretation of industrial mortality experience. *J Occup Med* **18**:169-170, 1976.
16. Enterline PE. Respiratory cancer among chromate workers. *J Occup Med* **16**:523-526, 1974.
17. Vital Statistics of the United States. Volume II — Mortality, Part A. Table 1-9 1-26 to 1-43, 1971.
18. Dement JM, Harris RL Jr., Symons MJ, and Shy C. Estimates of dose response for respiratory cancer among chrysotile asbestos textile workers. Presented at the Fifth International Symposium on Inhaled Particles, Cardiff, Wales, September 1980.
19. Enterline PE. Pitfalls in epidemiological research, and examination of the asbestos literature. *J Occup Med* **18**:150-156, 1976.
20. Milham S Jr. Methods in occupational mortality studies. *J Occup Med* **17**:581-585, 1975.